Participatory Technology Design: Active Learning Units

We want to prepare engineers of the future to attend to issues of justice in technology design. For this, it is important that students attend to social impacts of technology.

Technology designers need to learn how any technology design impacts people. User testing is one way that technology designers do this. The drawback is that user testing often starts after a prototype is already built. Design Thinking is another approach to engineering design that encourages human-centered design. Here, technology design begins with talking to people who the technology would benefit. Designers learn about people's experiences and how they want that experience to improve.

Participatory Design is yet a different approach. It assumes that people who are going to use a product are best situated to design key features of it. Everyday people also have knowledge relevant to technology design. Engineering students need to understand this. Users of reddit, google drive, Instagram, etc often request important design changes. These changes often improve the technologies.

The International Children's Digital Library is an example of participatory design. This was a project of Prof. Allison Druin at the University of Maryland. Prof. Druin and her team wanted to design a digital library for children. They involved children as designers in their project. This led to interesting outcomes. Child designers decided to organise the library in very atypical ways. Children preferred to organise the books by color, shape, and character, not by author names.

People from marginalised backgrounds are often not considered as partners in technology design. Yet, they are users of many technologies such as cell phones, social media, and e-commerce. The typical designs of technology can harm people from marginalized backgrounds. This includes people who are queer, transgender, working class, or disabled.

People from disadvantaged backgrounds have already influenced technology design in many ways. For example, seatbelts in cars were initially only tested for protecting male bodies. This made seatbelts less safe for women in the case of a car accident. It took feminist advocacy to change the design of car seatbelts to make them safer for women. Most web portal systems only allow a person to select "male" or "female" as their gender. This harms queer and transgender people. Advocacy from queer/trans groups has led to greater awareness. Now, some websites offer a wider range of options for choosing one's gender. Similarly, disabled people have designed more accessible systems and technologies. Disability justice advocates have shifted the landscape for making the world more accessible. Accessible websites, buildings, and sidewalks improve the experience of society at large.

Technology designers can often struggle to hear the perspectives of marginalized people. This results in technologies that can further harm marginalized people. To prevent this, instruction needs to bridge these social divide between designers and users of technologies. One way is to involve marginalized people as designers of technologies. We also need to build the motivation and commitment in future engineers to include marginalised folks in the design of new technologies.

Here, we present 3 activity-based modules that aim to addresses this point. These modules can be embedded in engineering ethics and engineering design courses. The first unit helps learners see that technology can also cause harm. The second unit is centered on an article on how the production of AI technologies causes harm to low-wage workers. The third and final unit presents learners with a case study encouraging them towards participatory design.

Unit 1: Technology can Cause Harm

The unit begins with introducing the idea that technology can cause harm. Then we provide some discussion questions. Students can reflect on the questions on their own or with a group of friends. Teachers can structure individual and collective reflection activities based on the reading and the discussion questions.

Technologies can cause harm

When we think about technologies and society, we often focus on how technology has helped make progress. Especially when we think about new technologies, we tend to focus on its potential benefits.

It is often harder for us to think about harms resulting from technologies. Say, that a person is physically harmed due to a device malfunction or an accident resulting from a negligence in the design of a device. That is something that we can understand and acknowledge. We talk about the engineers' responsibility in that case. But that is not the only kind of harm associated with technologies.

Technologies can impact mental health. We are seeing articles in newspapers that the very design of social media technologies is made such that people would be addicted to it. Social media use is also connected with depression and anxiety amongst people, especially youth. Social media can also shape our ideas and opinions. This can be used to manipulate people, as we saw in the Cambridge Analytica case during the 2016 US elections.

Take the search engine Google, for example. This has changed how we search and consume information online. However, relying on Google for most of our information has consequences. The search-engine algorithms can reinforce negative stereotypes about women, especially Black women. Similarly, Twitter algorithms can reinforce racism by preferentially showing white faces over darker faces.

Technologies for decision-making can cause harm

The police, courts, hospitals, and banks have started using Artificial Intelligence technologies in decision making. Sometimes, decisions from a computer seem more objective than those by humans. Computers and algorithms are seen as unbiased. However, the decision-making of most AI algorithms relies on statistical prediction. The AI programs develop ways of making these predictions by analyzing existing data. This ends up reproducing and exacerbating existing social inequalities. For example, say that an AI algorithm is training on public messages on a social media platform. If the posts within the data fed into the AI system has racist and sexist comments, the AI system might also produce racist and sexist output. If this output "feels" unbiased because it is generated by a computer, that can further harm people. This can allow racist actions, but with a coating of objectivity.

When AI technologies are used by the police and courts, their output could cause grievous harm! Increasingly, police departments are using AI based technologies to make decisions. But the data that is fed into these predictive policing systems can encode racial bias. This is leading to over-policing of Black neighborhoods and putting Black lives at risk in these neighborhoods.

Current and future engineers need to learn to acknowledge, recognize, and anticipate harm from new technologies.

Unit 1 Discussion Questions & Activities

1. What questions do you have after reading the above passages?

2. In the passages above, you read about the harms from technologies such as social media, Google search engine, and Artificial Intelligence technologies.

    1. Pick one topic from above, and try to find out more information about it online. Ironically, you might have to use the google search engine for the information. But you could also use an alternative such as Duck Duck Go. You can do this individually, or in a group of 3-4 friends/classmates.

    2. Make a list of who are the people who are harmed by the technology. What is the kind of harm they are facing from that specific technology?

3. As engineers, how should we orient to technology design and development, knowing that technology can harm people, especially people already marginalised?

Unit 1 Further Readings

1. Mayo Clinic Article on Social Media impact on adolescents: https://www.mayoclinic.org/healthy-lifestyle/tween-and-teen-health/in-depth/teens-and-social-media-use/art-20474437
2. Youtube Video: Sofiya Noble on racism embedded in google search: https://www.youtube.com/watch?v=iRVZozEEWIE
3. TED talk by Joy Buolamwini on facial recognition and racism: https://www.youtube.com/watch?v=UG_X_7g63rY
4. Will Douglas Heaven, “Predictive policing algorithms are racist. They need to be dismantled.” July 17, 2020, MIT Technological Review: https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/

Unit 2: Participatory Design and AI Ethics

On October 13, 2022, NEOMA magazine published an article titled, “The Exploited Labor Behind Artificial Intelligence.” The article was written by Adrienne Williams, Milagros Miceli And Timnit Gebru. In this unit, first some of the main points from the article are summarized. It would also help to read the article in full. The link to the article is inluded in the "Further Readings" section. The unit ends with some discussion questions which can be used for individual or group reflection activities.

Summary of the article

This article talks about the exploitation of workers in the development of Artificial Intellegence (AI) systems. Facial recognition, signature recognition, automated chat agents are examples of Artificial Intelligence technologies. AI systems aren’t really human-free. Humans are part of the systems. Here is how the authors say this:

> “Far from the sophisticated, sentient machines portrayed in media and pop culture, so-called AI systems are fueled by millions of underpaid workers around the world, performing repetitive tasks under precarious labor conditions.”

Who are these workers? and what are these repetitive tasks? To train AI systems, we need very large datasets. For example, ImageNET is a database of about 14 million images. These images would need to be processed for the AI systems. Often images or text that train an AI system need to be "cleaned" from content deemed offensive, such as pornography. If an AI system is trained on a dataset that includes racist or sexist expressions, then the system would also produce racist and sexist outputs. So the training data needs to remove such content. Similarly, the output of AI systems and algorithms need to be monitored and moderated to remove potentially offensive content. Data also need to be labelled with category labels that describe the nature of their content. Such tasks are performed by low-wage workers. People are paid very little, and have to work through very large amounts of data to earn a living. Here is how the authors describe this:

> “In 2009, however, Jia Deng and his collaborators released the ImageNet dataset, the largest labeled image dataset at the time, consisting of images scraped from the internet and labeled through Amazon’s newly introduced Mechanical Turk platform. Amazon Mechanical Turk, with the motto “artificial artificial intelligence,” popularized the phenomenon of “crowd work”: large volumes of time-consuming work broken down into smaller tasks that can quickly be completed by millions of people around the world. With the introduction of Mechanical Turk, intractable tasks were suddenly made feasible; for example, hand-labeling one million images could be automatically executed by a thousand anonymous people working in parallel, each labeling only a thousand images.”

And,

> “Content moderators, for example, are responsible for finding and flagging content deemed inappropriate for a given platform. Not only are they essential workers, without whom social media platforms would be completely unusable, their work flagging different types of content is also used to train automated systems aiming to flag texts and imagery containing hate speech, fake news, violence or other types of content that violates platforms’ policies. In spite of the crucial role that content moderators play in both keeping online communities safe and training AI systems, they are often paid miserable wages while working for tech giants and forced to perform traumatic tasks while being closely surveilled.”

While people are able to get jobs and earn a living doing this, this kind of work can harm their mental well-being:

> “Every murder, suicide, sexual assault or child abuse video that does not make it onto a platform has been viewed and flagged by a content moderator or an automated system trained by data most likely supplied by a content moderator. Employees performing these tasks suffer from anxiety, depression and post-traumatic stress disorder due to constant exposure to this horrific content.”

There is an international angle to how the companies operate. The authors note, "These tasks are often outsourced to precarious workers in countries like India, Kenya, the Philippines or Mexico."

The authors suggest a few ways forward. First, they want us to change our priorities. Folks interested in ethical AI systems have often focussed almost entirely on trying to make sure that their datasets are unbiased. The authors suggest that we shift the focus to "stopping the exploitation of labor in the AI industry."

But who should be figuring out how to do this? Often, we just assume that engineers, managers, CEOs, and such folks would be the ones to figure out solutions. Here, the authors are pushing us to consider another shift: the workers should be empowered to imagine and implement solutions towards their liberation from oppressive work conditions.

How should engineers help in this? By creating tools and work practices that help in shifting that power to the workers. For example, to do this work, we need more research on how labor practices are entangled with the development of AI systems. Labor practices shape the development of AI systems. But the development of AI systems also shapes labor practices. Practices such as crowdsourced labor have developed in parallel with AI technologies. The authors suggest that it would be crucial that workers are included in shaping the research agenda. And that they have the freedom to shift research purposes towards their needs. The research team should write or present the findings such that workers can understand them.

The article suggests that a variety of folks from different professions can support the workers. Journalists, artists, and scientists can help people understand workers' exploitation, how it happens, and build solidarity with their cause. And it is really important that technology workers in white collar jobs forge solidarity with the low-wage technology workers.

Unit 2 Discussion Questions & Activities

1. Why do you think AI companies are not simply paying the workers more for doing data processing and cleaning work? With higher wages a worker could work fewer hours, reducing the negative impact on them.

2. The authors note that, "We write this as people intimately connected to AI-related work. Adrienne is a former Amazon delivery driver and organizer who has experienced the harms of surveillance and unrealistic quotas established by automated systems. Milagros is a researcher who has worked closely with data workers, especially data annotators in Syria, Bulgaria and Argentina. And Timnit is a researcher who has faced retaliation for uncovering and communicating the harms of AI systems." *How do you think this informs the authors' perspectives?*

3. Find out a little more about each of the authors. Read up online about their past and present work. Where are they working? What are the main goals towards which they are working?

4. There are two shifts that the authors mention when they are discussing possible ways of moving forward. What are those shifts? What do you think about them?

5. ***A Mini Case Study:*** Say, you are working at an AI company working on generative AI technologies. Your company also draws on large datasets such as ImageNET. You are working in the Division of AI Ethics in the company. The head of the Division of AI Ethics happens to read the NEOMA article and wants to address this issue. She asks you to pull together a list of stakeholders who should be involved in discussions of possible solutions.
    (a) Who are the stakeholders in this issue? Read the article carefully to find out stakeholders. Make a list.
    (b) For each stakeholder write down how they might be benefiting and/or being harmed by this work. Again, the article would help you here.
    (c) How would you find out whether your ideas about stakeholders and the impact on them are accurate and comprehensive? Who/what might you be missing?
    (d) Who among these stakeholders would you invite to form the design team to search possible solutions or different ways of doing AI development work? Who would you leave out and why?
    (e) Have you included some people who are doing the kind of repetitive AI work on low wages that the article talks about? Why or why not?

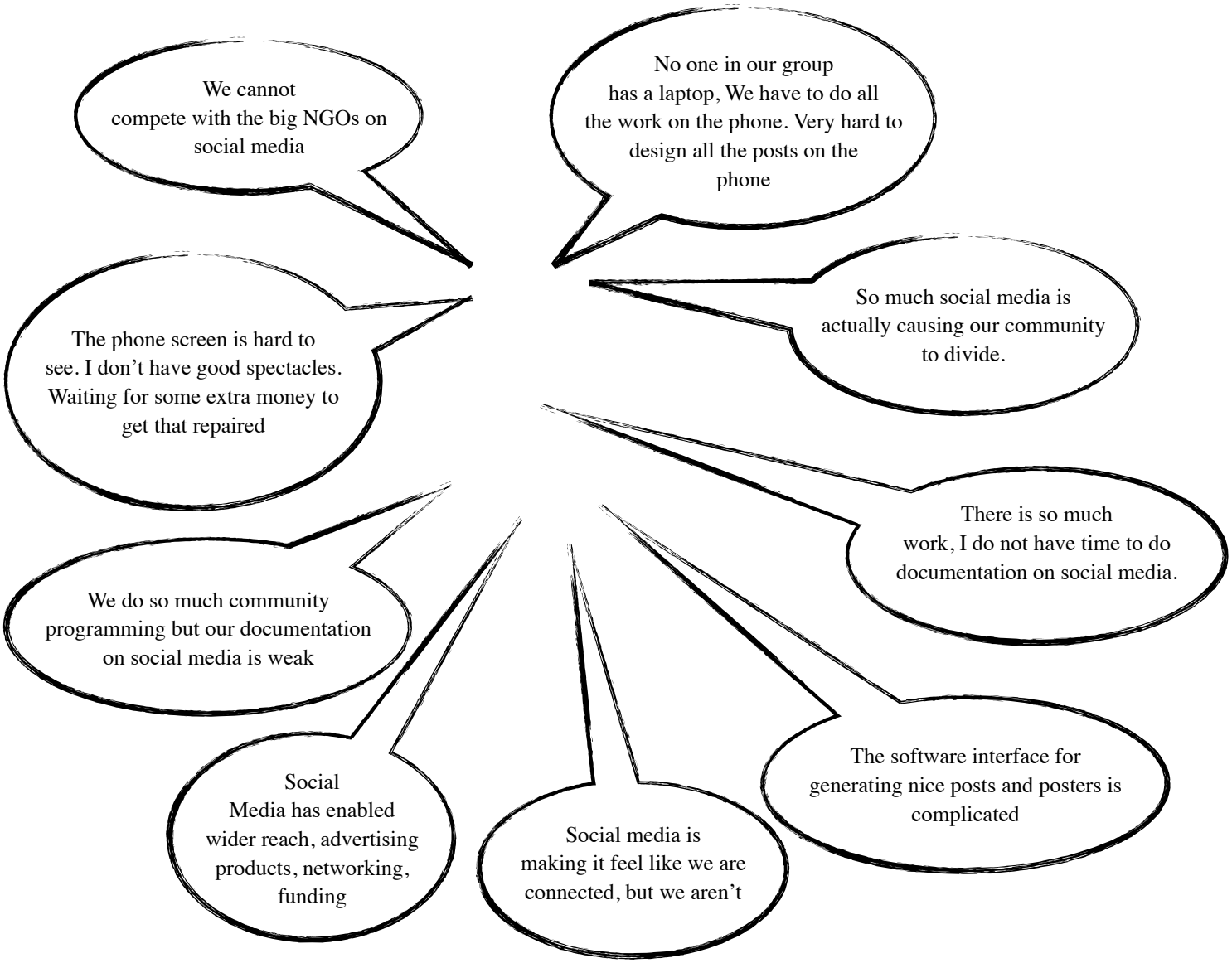
Unit 2 Further Reading

1. https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence/

2. https://www.image-net.org/about.php

Unit 3: Social media and community based organisations: Towards practicing participatory design

Say, you are part of a technology design team. Your team has been working with some community organisers on improving their experience of social media use. The community based organisations need to document their events and activities on social media, for outreach, for building community and membership, and for showcasing to funders. Over some visits, these are some of the comments that you have heard from the community organisers on social media use:


We cannot compete with the big NGOs on social media
No one in our group has a laptop, We have to do all the work on the phone. Very hard to design all the posts on the phone
The phone screen is hard to see. I don't have good spectacles. Waiting for some extra money to get that repaired
So much social media is actually causing our community to divide.
There is so much work, I do not have time to do documentation on social media.
We do so much community programming but our documentation on social media is weak
Social Media has enabled wider reach, advertising products, networking, funding
Social media is making it feel like we are connected, but we aren't
The software interface for generating nice posts and posters is complicated

Unit 3 Discussion Questions & Activities

3. Please discuss, in groups of 3-4 members each:

   (a) What are some of the main concerns pertaining to social media use for community organisers?

   (b) In what ways are they benefiting from social media use?

(c) Are there issues or comments that you do not fully understand? If so, discuss or ask an instructor.

(d) How would you want to proceed? What would you do next?

4. Did you (or your group) consider asking the community organisers to participate in the technology design process with you? If yes, what were your reasons for doing so? If not, why not? Discuss your responses with your group members. Write down some brief notes from your discussion.
5. Think of a few ways in which the community organisers could be involved as participants in the design process.

   (a) Discuss your responses with your group members.

   (b) Make a collective list.
6. What might be challenges your team encounters in working with the community organisers as partners in the design of the technology or the user experience?

   (a) Discuss your responses with your group members.

   (b) Make a collective list.
7. As a class, discuss how you feel about including the community organisers as design team members.

Unit 3 Further Reading

1. Hutchinson, H. B., Bederson, B. B., & Druin, A. (2006, June). The evolution of the international children's digital library searching and browsing interface. In *Proceedings of the 2006 conference on Interaction design and children* (pp. 105-112).
2. International Children's Digital Library: http://www.childrenslibrary.org/
3. Brown, T. (2008). Design thinking. *Harvard business review*, *86*(6), 84.
4. Sanders, E. B. N. (2002). From user-centered to participatory design approaches. In *Design and the social sciences* (pp. 18-25). CRC Press.
5. Papak, A., Gupta, A., & Turpen, C. A. (2018, June). Examining the relationships between how students construct stakeholders and the ways students conceptualize harm from engineering design. In *2018 ASEE Annual Conference & Exposition*.

Credits and Acknowledgements

These learning units have been designed by Ayush Gupta, Associate Professor, Dr. Homi Bhabha Centre for Science Education, Tata Institute of Fundamental Research, Mumbai. This work has been funded by PI Gupta's Infosys TIFR Leading Edge Grant.

We thank Gina Quan, Andrew Elby, David Tomblin, Nicole Mogul, Alexis Papak, and Jennifer Radoff for feedback on the learning units. We thank various grassroots non-profit organisations that provided help with the ethnographic components of this work.